The role of alternative salience in the derivation of scalar implicatures.

Alice Rees
Lewis Bott
School of Psychology, Cardiff University

Address for correspondence

Alice Rees
School of Psychology
Cardiff University
Tower Building
Park Place
Cardiff
CF10 3AT
Phone: +44(0)2920870470
E-mail: reesa34@cardiff.ac.uk

Abstract

Comprehension can be enriched by considering what a speaker could have said but did not; namely, the alternative. For example, “Betty passed some of her exams” can be interpreted as “Betty passed *some but not all* of her exams”. This enriched interpretation is an example of a scalar implicature. We consider whether the salience and use of the alternative are independent processes in the derivation of scalar implicatures or whether use is dependent on salience. Participants completed three sentence interpretation experiments in which the sentences invited scalar implicatures. The experiments used a structural priming paradigm with alternatives and implicatures as primes. We found that (1) adults could be primed to derive scalar implicatures when the alternative was the prime (2) they did so at a rate equal to if the scalar implicature itself were the prime. In the absence of evidence that the use of the alternative can be primed independently of its salience, we conclude that salience and use are not independent processes. Instead, we suggest that when the alternative is sufficiently salient, the implicature will automatically be derived.

Keywords: Implicatures, Alternatives, Structural priming, Pragmatics

The role of alternative salience in the derivation of scalar implicatures.

People often communicate much more than they explicitly say. For example, consider the following exchanges.

1. A: Are John and Mary coming to the party?
   B: John is.
   =>*Mary is not.*
2. I ate four doughnuts.
   => I ate *exactly* four doughnuts.
3. Betty passed some of her exams.
   => Betty passed some *but not all* of her exams.

In (1), B answers A's question about John coming to the party. Although B has not explicitly answered A's query about Mary's attendance, his utterance communicates that *Mary is not coming*. In (2), the listener can infer that the speaker ate *exactly* four doughnuts, even though the speaker did not explicitly say *exactly four*, and in (3), the listener can conclude that Betty passed *some but not all* of her exams, even though the speaker did not explicitly say *not all.*

Enrichments such as those above are commonly known as *scalar implicatures*. In each case the listener generated an enriched meaning based on the alternative to what the speaker said, that is, something that the speaker could have said but did not. There are many accounts of how implicatures can be derived but most assume something like the following, inspired by Grice (1975): (i) The listener computes the basic meaning of an utterance, (ii) recognises that an alternative phrase could have been used, (iii) negates the alternative and (iv) combines this with the basic meaning. For example, in (1), Speaker A recognises that B could have said "John and Mary are coming to the party" (the *alternative*). Since B did not say this, and assuming that she is being cooperative, A can infer that "John and Mary are coming to the party" is not true. Thus, combining what is said, *John is coming to the party*, with the negation of the alternative, *it is not the case that John and Mary are coming to the party*, the listener arrives at the meaning that *John but not Mary is coming to the party*. Similar reasoning can be used to derive the enrichment seen in the other examples. In (2), since the speaker said *four* but not *five, six, seven, etc*., the listener can infer that *not five, not six, not seven* is the case, and conclude that the speaker means *four but no more*. In (3) the speaker could have said *all*, but since they did not, the speaker can infer *not all*.

Implicatures are optional: the listener chooses whether to incorporate an implicature into the sentence meaning. For example, in (3), if the preceding discourse had been about whether Betty would pass *any* of her exams, the listener would likely not derive the *not all* inference (since the *not all* part would be largely irrelevant). Understanding how and why certain contexts cause people to enrich the basic meaning of expressions has been a fundamental research goal in pragmatics (e.g. Chierchia, 2013; Geurts, 2010; Grice, 1989; Horn, 1972; 1989; Levinson, 2000) and psycholinguistics (e.g. Bott, Bailey, & Grodner, 2012; Bott & Chemla, 2016; Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006; Breheny, Ferguson & Katsos, 2013; Degen & Tanenhaus, 2015; Gotzner, Wartenburger, & Spalek, 2016; Grodner, Klein, Carbary, & Tanenhaus, 2010; Huang & Snedeker, 2009a; Tomlinson, Bailey & Bott, 2013). In our study we address the role of the alternative in this process. We test whether the salience of the alternative entirely determines whether an expression will be enriched, or whether an additional, independent usage mechanism is justified.

Combination and salience models of implicature

Most researchers agree that there are two stages to the implicature process. The first is that a relevant alternative is retrieved from the lexicon or the context, or constructed. The second is that this alternative is negated and combined with the basic meaning of the sentence. However, it is not clear how the second stage depends on the first. The second stage could apply automatically once the first stage is complete, so that the implicature is always derived if the alternative is sufficiently salient, or the second stage could be activated independently of the first. We refer to the former possibility as the *salience model*, since the implicature depends purely on the salience of the alternative, and the latter as the *combination model*, since the implicature depends on a combination of the salience of the alternative and the activation of an independent, usage mechanism (see Figure 1).

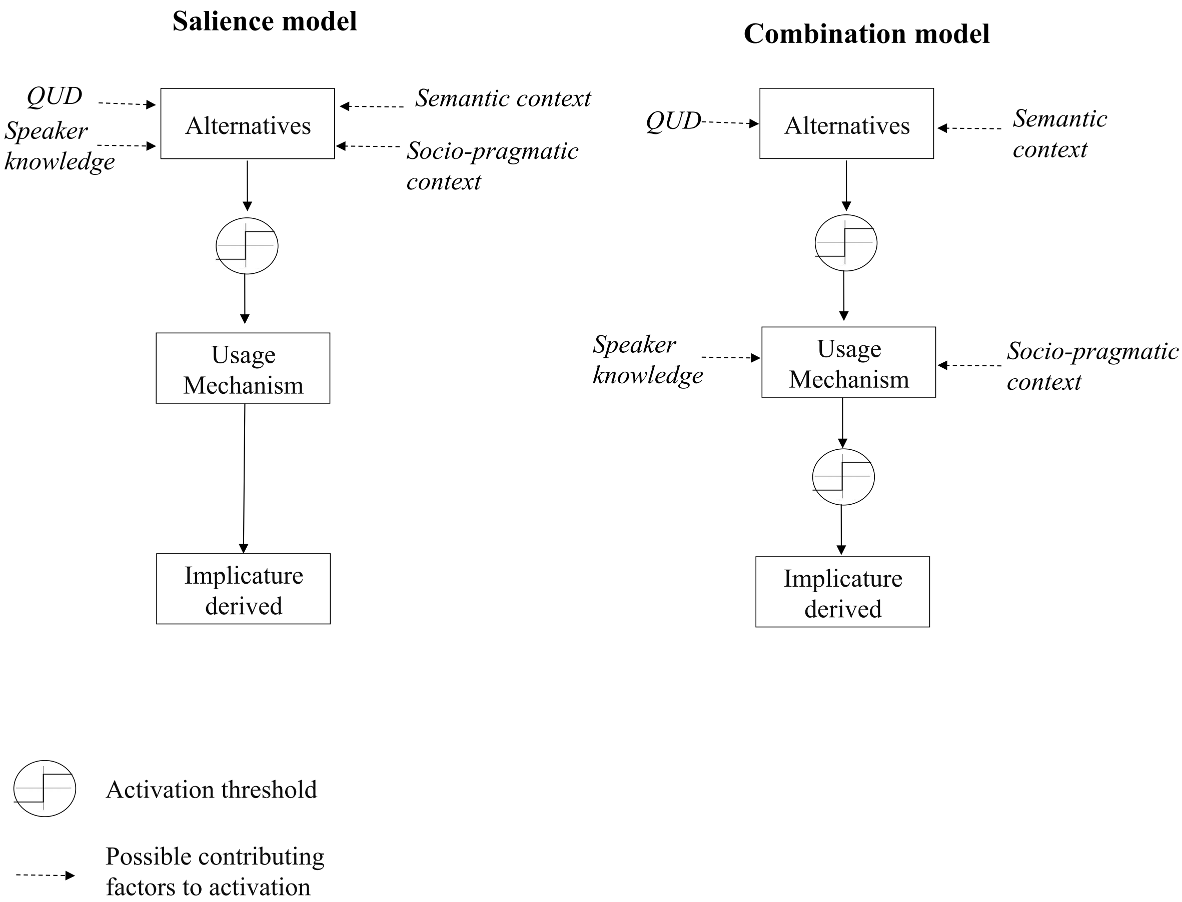

Figure 1. Salience and combination models. Alternatives have varying levels of activation. For the salience model, the usage mechanism is automatically applied after the alternatives obtain sufficient activation to exceed a threshold but for the combination model, the alternatives and the usage mechanism are required to exceed a threshold. Note that both models have the same number of processing steps but the combination model assumes the usage mechanism can be independently activated whereas the salience model does not.

Both models assume that alternatives have varying degrees of activation. For the salience model, if the activation of any one alternative exceeds a threshold, the usage mechanism will be applied, and the implicature will be computed. The usage mechanism is not modulated independently of the activation of the alternatives. The combination model also assumes that the alternatives have varying degrees of activation and a threshold, but additionally assumes that the usage mechanism does. For the salience model, contextual factors (e.g., the question under discussion, whether alternatives have been mentioned in the discourse, speaker knowledge) affect the activation levels of the alternatives, but for the combination model, contextual factors affect activation levels of the alternatives and/or the usage mechanism, independently. Both models explain how the implicature arise in some circumstances but not others. For example, in (1) the salience model explains the implicature by assuming that the alternative (*John and Mary are going to the party*) is sufficiently active that it exceeds the threshold necessary to trigger the usage mechanism and so generate the implicature. The combination model also requires alternatives to be sufficiently active but additionally assumes that the usage mechanism is active. Similarly, consider a situation where an implicature would not arise. In (1), assume that B knows information about John but knows nothing about Mary, and the listener is aware of this. Under these conditions the *not Mary* implicature does not arise (the competency assumption; see Grice, 1989, and Sauerland, 2004). The salience model explains the absence of implicatures by assuming that the absence of speaker knowledge suppresses activation levels of the alternatives to such a degree that the usage mechanism is not triggered. The combination model explains this by assuming either that the alternatives are not sufficiently active, or that speaker knowledge directly suppresses the usage mechanism. The crucial difference between the two is that with the independent usage mechanism, the combination model has an extra method of accommodating contextual factors, such as speaker knowledge or alternative relevance.

Implicature theories from the formal pragmatics literature can be broadly mapped on to the salience/combination distinction. Among the salience models are Grice's original account and its more recent developments, the Neo-Gricean models (Horn, 1972; Levinson, 2000). Grice's account assumes that if there is linguistic material that is relevant and more informative than the basic expression, this material should be designated an alternative. Subsequently the alternative is combined with the basic meaning of the sentence using domain general reasoning processes. Because there is no mechanism for blocking the implicature from arising after the alternative has been made active, the usage mechanism should not be thought of as independent from the activation of the alternative. Thus Grice's model can be classified as a salience model. While Neo-Gricean accounts provide more detail on how the alternative is formed, the basic assumptions regarding the automatic application of the usage mechanism are the same. There is no mechanism proposed for blocking the implicature after the alternative has been identified and so the usage mechanism cannot be considered independent. This seems particularly clear with regards to Levinson, who presents extensive arguments in favour of implicatures being the result of default reasoning processes. If a usage mechanism were able to block the implicature from arising in certain contexts, the system would no longer be a default model. Recent grammatical models of implicatures should also be considered salience models for the same reason. Indeed, Chierchia et al. (2012) explicitly identify with salience model, "providing alternatives are active, such alternatives are obligatorily factored into meaning… if the alternatives are not active the plain unenriched meaning is used and no scalar implicature comes about." (p.2304).

Other formal accounts are more readily mapped onto combination models. These typically assume an independent usage mechanism to explain the effects of relevance or speaker knowledge. Sauerland (2004) and the structural theory of alternatives (Katzir, 2007; Fox & Katzir, 2011) could sensibly be implemented in this way. According to Sauerland, individuals first compute a primary implicature, in which the speaker is assumed to be uncertain about whether a stronger statement holds. Under the right epistemic conditions, the primary implicature is then strengthened to a secondary implicature. An independent usage mechanism could be linked to the assessment of epistemic context so that it is only activated when the context is appropriate. Structural theories of alternatives also assume multiple steps in the derivation process. For Fox and Katzir (2011), the set of alternatives is first determined by structural linguistic or discourse factors (e.g. sentence complexity) but at a later stage, conversational relevance filters out inappropriate alternatives. As with Sauerland, an obvious implementation of this would be for relevance to affect the probability of the usage mechanism being activated. More generally, researchers often describe the implicature process as a series of discrete stages (e.g. Breheny, Ferguson, & Katsos, 2013; Geurts, 2010; Katsos & Bishop, 2011), with the usage mechanism described as a separate from the identification of alternatives. For example, Katsos and Bishop (2011) say, "The first step involves determining whether the speaker could have made a more informative statement….The second step involves the negation of the more informative statement that was identified in the first step." (p. 68). Similarly, Breheny, Ferguson, and Katsos include speaker knowledge as a separate step in implicature generation (Table 1, p.424). While we doubt these researchers are committed to a combination model, the exposition implies some degree of independence between the formation of the alternatives and their negation, such that the usage mechanism might be affected by factors other than the salience of the alternatives.

The forgoing discussion demonstrates that both salience and combination models are plausible implementations of formal pragmatic models. There is no reason to choose one model over the other on the basis of a consensus in the formal literature. However, the models discussed above were not intended as mechanistic, psychological accounts and in many cases the question of whether there is an independent usage mechanisms depends more on implementation than the structural properties of the theory. For example, we classed Fox and Katzir (2011) as a combination model (partly on the recommendation of an anonymous reviewer), but conversational relevance might act to prune the set of alternatives as part of the initial stage of alternative generation, rather than the later stage of alternative usage. Consequently Fox and Katzir could still be implemented in a system where the usage mechanism was automatically activated on the basis the activation levels of the alternatives. We therefore emphasise that we are not testing any particular formal model but the broad implementation principles behind the salience/combination distinction.

Alternatives in scalar implicatures

There have been no previous attempts to test between the salience and combination models but there is experimental evidence of the importance of the alternative more generally. Much of this comes from the developmental literature. Developmental data suggest that children have difficulty generating scalar implicatures (e.g. Huang & Snedeker, 2009b; Noveck, 2001; Papafragou & Musolino, 2003) and there is now converging evidence that at least part of the reason is children's difficulty in generating appropriate alternatives (Barner, Brooks, & Bale, 2011; Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Gualmini et al., 2001; Skordos & Papafragou, 2016).

A particularly convincing example of this is Skordos and Papafragou (2016). They tested whether the accessibility and relevance of alternatives affected five-year-old's ability to generate scalar implicatures. Children completed a sentence-picture verification task where children were presented with a picture and had to judge the felicity of the description provided. The critical sentences were underinformative statements involving *some*. Accessibility of the alternative was manipulated by varying the order of trials. In the *mixed* condition trials using *some* and *all* were intermixed so that *all* was accessible during the evaluation of *some*. In the *some-first* condition all of the trials using *some* were presented before trials using *all.* When the alternative, *all*, was made more accessible to children they derived more scalar implicatures than when the alternatives were not accessible. Consistent with Barner, Brooks, and Bale (2011), this result demonstrates that the accessibility of the alternative partly determines whether children derive the implicature. Skordos and Papafragou go further than demonstrating that accessiblity is important, however. In later experiments they show that the alternative needs to be relevant, as well as accessible. Relevance was determined by the sentence evaluation criteria. In the *all* relevant condition, the task required children to process the quantifier to determine whether the sentence was true, whereas in the *all* irrelevant condition, they heard the same sentences but the composition of the picture meant that the quantifier was not relevant (the truth of the sentence depended on only the predicate). Skordos and Papafragou observed higher rates of the implicature in the *all* relevant condition than the *all* irrelevant condition, even though *all* was accessible in both cases. A further experiment revealed that an equally high rate of implicatures was derived when a relevant *none* was used instead of *all*.

Skordos and Papafragou (2016) provide an important demonstration of how children are influenced by the salience of the alternative. Their results are consistent with the salience and the combination model, however. The salience model could explain the results by claiming that *all* was made more active in conditions where there was a high rate of implicatures. In Experiment 1, repeated processing of *all* in the *all* condition raised activation levels; in Experiment 2, depth of processing for *all* was greater in the relevant condition than the irrelevant condition, and activation levels were greater as a consequence; and in Experiment 3, the semantic association between *none* and *all* raised activation levels of *all* (more generally, the set of quantifiers was activated when the question-under-discussion was recognised as being about quantifiers).

While there is general agreement that the salience of the alternatives is important in the adult processing literature, alternatives have received far less direct interest than in the developmental literature. Indirect evidence of the role of the alternative salience is given by Chemla and Bott (2014). They tested whether free choice inferences were a form of complex scalar implicatures using a sentence verification task. In one block participants completed a free choice task where they were presented with a scenario and a sentence using disjunction and had to judge the truth of the sentence based on their knowledge and the scenario. The veracity of the sentence depended upon whether or not participants derived a free choice inference. In the other block participants completed a similar task but using scalar implicatures rather than free choice inferences. Chemla and Bott found that while scalar implicatures were delayed relative to literal controls (replicating Bott & Noveck, 2004), free choice inferences were not. They concluded that either free choice inferences were a completely different phenomenon from scalar implicatures, or that the alternatives used to derive free choice inferences were sufficiently salient to the participant that retrieving them was not a costly process. Under the latter explanation, the cost observed when people derive scalar implicatures can at least partially be explained by the inaccessibility of the alternative. If the alternatives are salient, implicatures can be computed quickly, whereas if the alternatives are not as salient, implicatures will cause a processing delay (see also van Tiel and Schaeken, 2016). While this theory relates to processing cost, and not implicature rates, it nonetheless suggests that alternative salience is important for adult processing of implicatures.

Bott and Chemla (2016) also argued that the salience of the alternative was important. They tested whether people could be primed to derive scalar implicatures, in the same way that they can be primed to produce particular syntactic structures (e.g., Bock, 1986; Bock, Dell, Chang, & Onishi, 2007; Branigan, Pickering, & Cleland, 2000; Thothathiri & Snedeker, 2008; see Pickering & Ferreira, 2008 for a review). Participants saw prime trials, in which they derived either an implicature interpretation (a strong interpretation) or a non-implicature interpretation (a weak interpretation), and target trials in which the sentence was ambiguous between a strong and a weak interpretation. Target trials appeared subsequent to prime trials. They found that the prime sentence influenced the interpretation assigned to the target sentence such that a strong prime trial led to more strong target interpretations than a weak prime trial. They proposed several explanations for the priming effect but one involved the strong prime making the alternative more salient to the participant, which in turn elevated the rate of implicatures.

In summary, there is a range of evidence suggesting that the salience of alternatives is important for deriving scalar implicatures but little direct evidence to distinguish between the salience model and the combination model. In our study we provide a direct test between them. The logic is explained in the next section.

Retaining activation levels across time

The salience and the combination models are not easy to distinguish empirically. Variable activation levels of the alternatives can explain many findings, and these are present in both models. To test between them then we make a further assumption about the usage mechanism. We suggest that just as the alternatives maintain activation levels across time, the usage mechanism also maintains activation levels across time. This means that if the usage mechanism has been used earlier in the discourse, activation levels should remain high, and if it has been suppressed, they should remain low. The rationale for this assumption comes from research on structural priming (see Pickering & Ferreira, 2008; and Branigan & Pickering, 2017, for reviews). These studies demonstrate that many linguistic structures maintain activation levels in just this way (e.g. active and passive syntactic structures, Bencini & Valian, 2008; transitive and dative syntactic forms, Bock 1986; animacy assignments, Bock, Loebell, & Morey, 1992; conceptual level structures for configurations in a maze, Garrod & Anderson, 1987; scopal interpretations with "every", Raffray & Pickering, 2010, and Chemla & Bott, 2014). For example, Branigan and Pickering (1998) showed that when a confederate used a double object structure to describe a picture ("give the man the book"), participants were more likely use a double object structure in subsequent trials than a prepositional structure ("give the book to the man"). Activation levels of the double object structure were retained across time so that when the participant needed to choose a structure, the double object structure was more active than the prepositional structure.

The consequence of assuming that activation levels are maintained across time and linguistic space is that the combination model now predicts the usage mechanism can be primed by recent use. If the usage mechanism has been used recently in the discourse, it should have higher activation levels than if it has not. Since higher activation levels translate as a greater probability that the usage mechanism will be triggered (see Figure 1), the more recent the application of the usage mechanism, the more likely the usage mechanism is to be applied to the current interpretation. The basic logic of our design follows that of researchers in structural priming who claim that the presence of a priming effect reflects the presence of a structure, and conversely, the absence of a priming effect reflects the absence of a structure (Branigan & Pickering, 2017; Pickering & Ferreira, 2008).

In our study we adapted Bott and Chemla's (2016) paradigm to test whether the usage mechanism can be primed. A priming effect would support the combination model, and the absence of one would support the salience model.

Current Study

Participants completed a picture-sentence matching task. They were presented with two pictures and a sentence, and had to select which of the two best matched the sentence. Participants' interpretation of the sentence was indicated by their picture selection. There were prime trials and target trials. Target trials were presented immediately after prime trials. The correct interpretation of the sentence in prime trials was unambiguous, but in target trials participants could choose between an implicature reading (a *strong* interpretation) and a non-implicature reading (a *weak* interpretation). A priming effect was shown when the type of prime trial influenced the interpretation of the target trial.

There were three types of prime: strong, weak, and alternative. In strong and weak prime trials, sentences contained a scalar trigger term, e.g. "some" (see Figure 2). In alternative prime trials, sentences contained a more informative alternative to the scalar term, e.g. "all". In strong trials, the picture configuration meant that the most relevant interpretation was a strong reading (implicature) e.g. *some but not all.* In weak trials the relevant interpretation was a weak reading (non-implicature) e.g. *some and possibly all*, and in alternative prime trials, the relevant interpretation was the alternative reading, e.g., *all.*

Target trials consisted of a picture corresponding to the weak interpretation of the sentence, and a picture with the words, "Better Picture?" (modelled on the "hidden box" paradigm of Huang, Spelke, & Snedeker, 2013; see Figure 2). If participants felt that there was a picture that better matched the sentence than the one shown, they could select the "Better Picture" option. Thus, weak interpretations of the sentence were measured by selecting the weak picture and strong interpretations by selecting the "Better Picture" option. Priming of scalar implicatures occurred when there were more strong target interpretations subsequent to a strong prime than a weak prime.

Three categories of expression were used: quantifiers, numerals, and *ad hoc* constructions. All three are argued by some authors to be types of scalar implicature (e.g., Hirschberg, 1991; Horn, 1972, 1989; van Rooij & Schulz, 2006). In particular, sentences using these expressions admit strong and weak readings and the strong reading can be derived using a Gricean reasoning process. For example, the weak interpretation of "There is a V" is *there is a V and possibly other letters*, and a relevant alternative (in the context presented in Figure 2) is "There is a V and a C." Combining the negation of the alternative, *there is not a V and a C,* with the basic meaning of the sentence, *there is a V and possibly other letters*, yields the strong meaning, *there is a V and nothing else*. A similar reasoning process applies to the numbers. We used a range of scalar implicature expressions, rather than the prototypical *some*, because we wanted to test whether our findings applied to a range of scalar trigger expressions.

According to both the salience and the combination models, alternative and strong prime trials should increase the salience of the alternative relative to the weak prime trials. In the alternative trials, the alternative is made explicitly available, and in the strong trials, the alternative is made implicitly available by forcing the participant to derive a scalar implicature. Therefore the rate of implicature in target trials should be higher following alternative and strong prime trials than weak trials. The models make different predictions for the priming caused by the strong relative to the alternative prime, however. The salience model assumes that the rate of implicature is entirely dependent upon the activation of the alternative. Since the alternative prime makes the alternative active, and the strong prime also makes the alternative active (via the scalar implicature), the alternative and the strong prime should lead to equal rates of implicature. In contrast, the combination model assumes that activation of alternatives is only one component of what determines whether an implicature is derived. The other is activation of the usage mechanism. Assuming the usage mechanism maintains its activation levels across trials (as described above), it should continue to be active after a strong prime (which triggers the usage mechanism), but not an alternative prime (which does not). For the combination model then, the alternative prime should raise the activation of the alternative, leading to a certain amount of priming, but the strong prime should raise the activation of the alternative and of the usage mechanism, subsequently leading to more priming than the alternative alone. Thus the combination model predicts greater priming following the strong prime trials than the alternative prime trials.

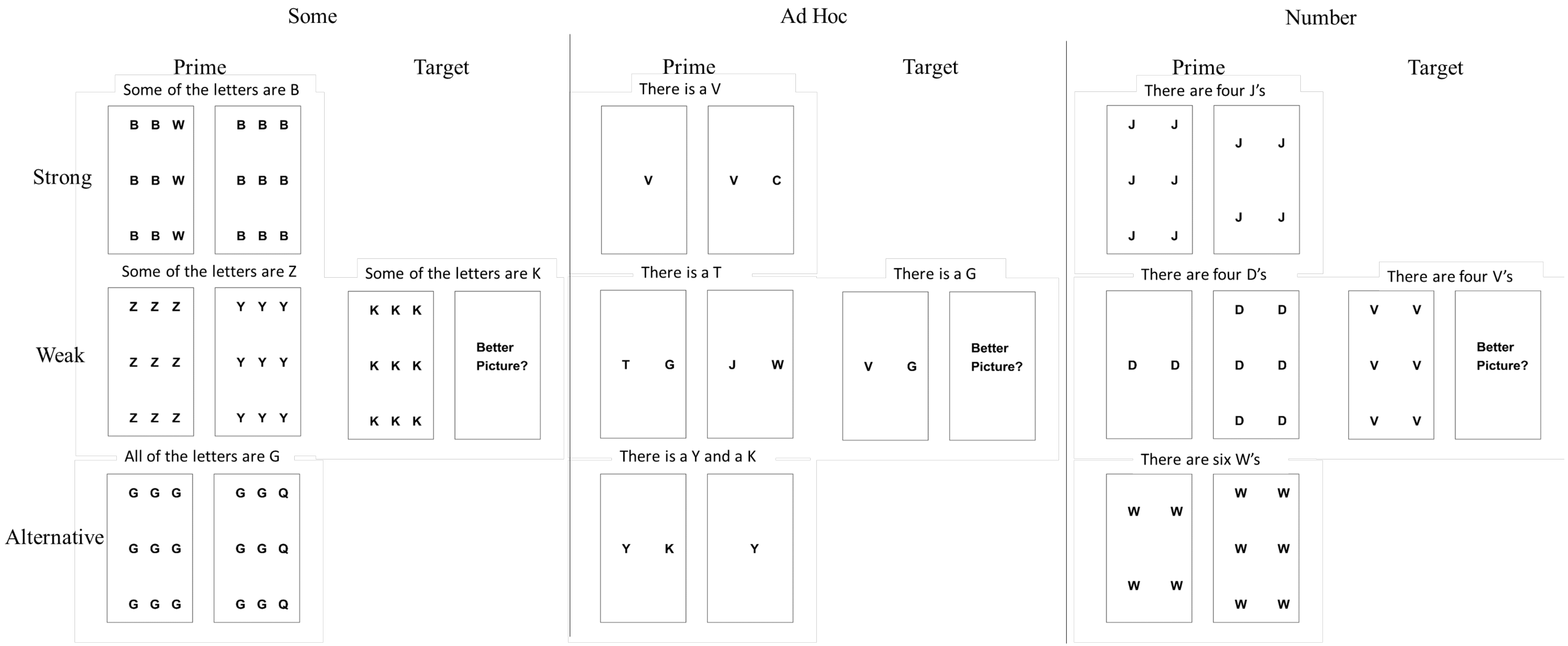

Figure 2. Example primes and target for each implicature category. Left to right: *some*, *ad hoc*, and number. Primes top to bottom: strong, weak, and alternative. Each trial consisted of a sentence and two pictures. Participants selected the picture that best matched the sentence e.g. for the *some* alternative prime trial (bottom left), "All of the letters are G", participants would select the left hand picture.

Experiment 1

Method

Participants

One-hundred Cardiff University undergraduate students participated for course credit (93 female, average age 19.4 years). All participants were native English speakers.

Materials

Each trial consisted of a sentence presented above two pictures. Participants had to select which one of the two pictures best matched the sentence. Pictures consisted of rectangles containing either letters or the text "Better Picture?"

Three expression categories were used: *some*, number, and *ad hoc*. The expressions have a weak meaning that can be enriched to form the strong meaning. Table 1 shows example expressions, together with plausible alternatives and the subsequent implicature.

Implicature category	Expression	Alternative	Implicature
Some	Some	All	Some but not all
Number	There are N	N+1,+2…	N exactly
Ad hoc	There is an X	X and Y	There is an X and nothing more

Table 1. Experimental stimuli.

Prime trials consisted of two pictures with letters inside. There were one, two, four, six, or nine letters depending upon the implicature category of the prime. Pictures with letters could either be strong, weak, or false depending upon the predicate of the sentence. For each implicature category there were two possible sentence frames (see below). One was used in the strong and the weak primes and the other was used in the alternative prime trials. Weak prime trials involved a weak and a false picture, and strong primes involved a strong and a weak picture. Alternative primes had the same picture configuration as strong primes, but due to the different sentence frame used, alternative primes had a weak and a false picture.

Some prime trials. For *some* prime trials the sentence frames were "Some of the letters are [letter]" for strong and weak primes and "All of the letters are [letter]" for alternative prime trials. The two pictures in *some* prime trials contained nine letters. These pictures could be strong, weak, or false depending on the sentence predicate. The nine letters in strong pictures were made up of six letters which matched the predicate and three other letters. Weak pictures contained nine letters that matched the predicate. False pictures also contained nine but these were different from the predicate.

Number prime trials. Strong and weak prime trials used the sentence frame "There are four [letter]'s". Alternative trials used "There are six [letter]'s". Strong pictures contained four letters, weak pictures contained six letters, and false pictures contained two letters. All of the letters matched the predicate.

Ad hoc prime trials. The sentence frame in strong and weak prime trials was "There is a [letter]". The alternative trials used "There is a [letter] and a [letter]". Pictures contained either one or two letters. Strong pictures contained a single letter. Weak and false pictures contained two letters. The letter in the strong picture was the same as the predicate. Weak pictures contained one letter that was the same as the predicate and one that was different. Both of the letters in false pictures were different from the predicate.

Target trials. Target trials also contained two pictures. One contained letters and was a weak picture. The other was a "Better Picture" option, similar to the covered box paradigm (e.g. Huang et al., 2013). The sentence frame in target trials was the same as the one used in strong and weak prime trials. Consequently, if participants selected the picture with letters this was consistent with a weak (non-implicature) interpretation of the sentence. Selecting the "Better Picture" option corresponded to a strong (implicature) interpretation.

Design

There were three types of expression (*some*, number, *ad hoc*). For each there were three types of prime (strong, weak, alternative). Consequently there were 3 (expression) x 3 (prime) = 9 conditions.

There were two primes for every target. This was done to boost the effect of the prime. Thus an experimental item was a triplet of two primes and a target. In prime trials the position of the correct picture was systematically varied (left or right) to prevent participants from becoming biased to pictures in a particular position. The positions were crossed so that there were 4 patterns of correct responses across the two prime trials (left-left, right-right, left-right, or right-left). Consequently each condition had four prime combinations. This resulted in 4 (combinations) x 9 (conditions) x 3 (triplets) = 108 experimental trials. Triplet presentation was randomised so that the order was different for each participant.

Thirty-six single filler trials were included. Filler trials were a mixture of prime and target trials which were indistinguishable from their experimental counterparts. The only difference between filler and experimental target trials was that instead of a weak picture the filler trial had a false picture. The false picture had the same configuration as a weak picture but the letters were inconsistent with the sentence predicate. The filler target trials would therefore give participants the opportunity to select "Better Picture" and thus consider this as an acceptable response.

Procedure

The experiment was run as an online questionnaire using Qualtrics Survey Software. Participants were told to "select the picture which best matches the sentence" and to select the "Better Picture" option if the picture did not match the sentence. Participants were shown examples of prime and target trials with the correct image selected. The examples were accompanied by an explanation of why the selection was correct.

Participants responded by clicking a box under the appropriate picture with the mouse. Participants then clicked the "next" button to show the next trial.

Results

Analysis Procedure

Responses to target trials were removed if the corresponding prime trial was not answered correctly. In Experiment 1, 4% of the data was removed for this reason. Consequently, during pairwise comparisons these participants were not included in the analysis. The remaining data underwent a logit transformation.

Data were analysed using 3x3 ANOVA with prime type (strong, weak, and alternative) and expression (*some*, *ad hoc*, and number) as within-subjects factors. We used Bayes factors to interpret the nonsignificant findings (Dienes, 2011, 2014; Rouder, Speckman, Sun, Morey, & Iverson, 2009). We used the default JZS prior (Rouder et al., 2009) for all analyses. The JZS prior is a non-informative objective prior that minimises assumptions regarding expected effect size. Bayes factors using the JZS prior (0.707) were calculated using JASP (JASP Team, 2016). Bayes factors > 3 suggest 'substantial' evidence for the alternative hypothesis and Bayes factors < 0.33 indicate 'substantial' evidence for the null hypothesis (Dienes, 2011, 2014).

Analysis and Discussion

Figure 3 shows the rate of implicature to targets as a function of prime and expression. Three patterns are noteworthy. First, the overall rate of implicature varied across expressions, as shown by a main effect of expression, $F(2, 186) = 51.64$, $p < .001$. For *ad hoc* expressions, there was a clear bias towards the weak interpretation, whereas for *some* and the numbers the split was more even (although still biased to towards weak interpretations). This pattern corresponds with our intuitions, in that the weak interpretation seems particularly salient for the *ad hoc* expressions, but it is not obvious what causes this result. We return to explanations of this effect in the General Discussion.

Second, rates of strong interpretations were higher following strong and alternative prime trials than weak primes, as shown by the main effect of prime, $F(2, 186) = 32.24$, $p < .001$. Planned comparisons showed that rates of implicature were significantly higher following strong primes than weak primes, $t(99) = 6.67$, $p < .001$, illustrating the basic priming effect observed by Bott and Chemla (2016). There was also significantly greater rates of implicature following alternative primes than weak primes, $t(99) = 6.69$, $p < .001$.

Finally, there was no difference between rate of implicature following strong and alternative primes, $t(99) = .13$, $p =.89$, BF = 0.11. That the BF was less than 0.33 suggests strong support for the null hypothesis, and consequently, support for the salience model.

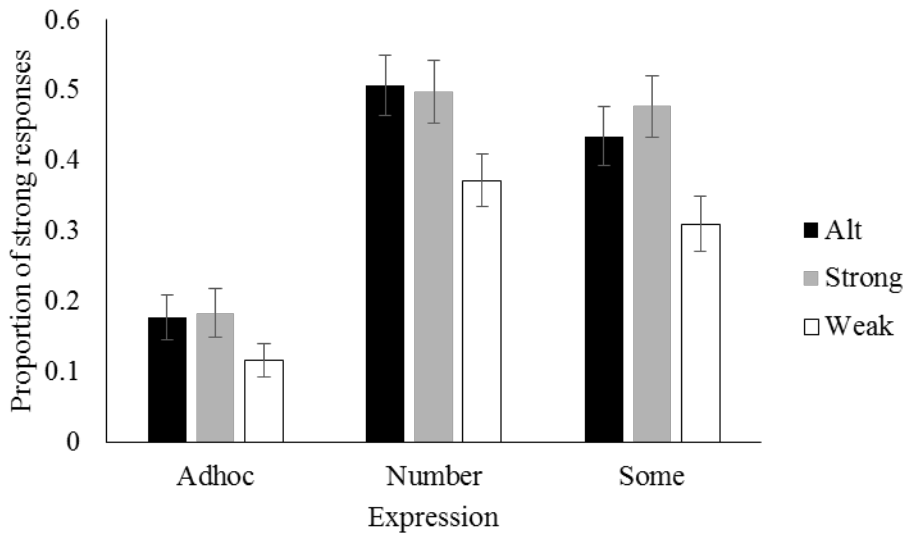

Figure 3. Proportion of strong (implicature) responses to target trials by expression. Error bars show standard error.

While there was no significant interaction between prime and expression $F(4, 372) = 2.00$, $p = .094$, BF = 1.84), the BF suggests that the experiment was insensitive in this respect. We therefore considered each expression separately. As Figure 3 shows, the same pattern was found across each. There were significantly higher rates of implicature following strong and alternative primes than weak primes for *ad hoc,* number, and *some* expressions, $t(95)$'s > 2.69, p's < .009; $t(99)$'s > 4.38, p's < .001; $t(97)$'s > 4.51, p's < .001, respectively, and no difference between strong and alternative primes across expressions, $t(95) = .44$, $p = .66$, BF = 0.12; $t(99) = .20$, $p = .84$, BF = 0.11; $t(97) = .94$, $p = .35$, BF = 0.17. In short, the data suggest that the alternative prime was just as effective as the strong prime, regardless of whether the expressions were considered separately or as a whole.

Individual participant details

It is possible that the priming effects reported above were due to participants changing interpretation only once during the experiment. Perhaps a participant started the experiment with a particular interpretation, strong or weak, but then switched to a different one after they realised that in this particular context the opposite interpretation was more appropriate. They then maintained the alternative interpretation throughout the experiment. This would be a global priming effect linked to particular experimental contexts, rather than a local priming effect. To investigate this we examined how often individual participants switched from one interpretation to another across the experiment. Figure 4 shows the results.

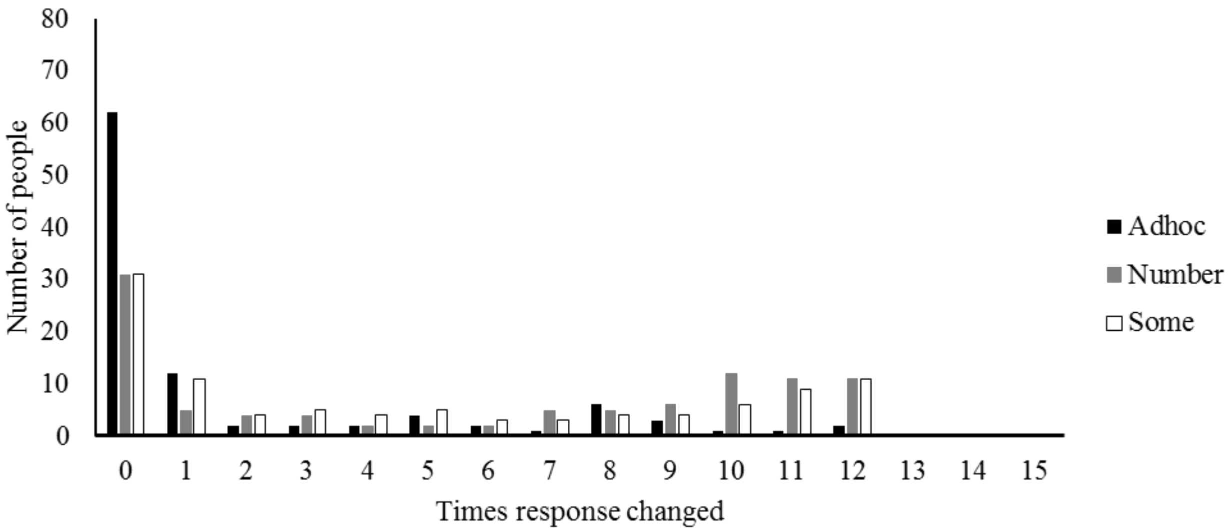

Figure 4. Frequency of response changes.

Consider first the *some* expressions. Out of 100 hundred participants, 31% maintained the same interpretation throughout the experiment (either strong or weak). A further 11% changed only once, but the remaining 68% changed at least twice, such as from strong to weak and back again to strong. The results for the numbers are similar. Around a third could not be primed at all, but around two thirds changed interpretation at least twice. There were more participants who could not be primed for the *ad hoc* expressions, 62%, and 12% changed only once, but nonetheless 26% changed at least twice. The average priming effects consequently cannot be due to participants changing their interpretations once only.

Summary

There were significantly more strong interpretations after a strong prime than a weak prime, replicating Bott and Chemla (2016), and significantly more after the alternative than the weak prime. Crucially, the rate of strong interpretations did not differ between the strong and alternative interpretations, and the Bayes Factor analysis suggests that this was not because the experiment was insensitive to reasonable effect sizes. Consequently the result supports the salience model.

The findings from Experiment 1 suggest that alternative and strong prime trials raised the activation of the alternative so that on subsequent target trials, participants were more likely to derive scalar implicatures. In Experiment 2 we test a different explanation based on the visual structure of the pictures.

Experiment 2

The strong and alternative primes in Experiment 1 exposed participants to a strong picture. For example, the strong *some* prime and the alternative *some* prime exposed participants to a picture with nine letters, three of which were different to the others (*some but not all*). In contrast, the weak prime did not expose participants to a strong picture. Thus, a possible explanation for our findings (and those of Bott & Chemla, 2016), is that exposure to the strong picture caused participants to reject the weak option in the target trial. Perhaps participants became aware that a strong interpretation of the target sentence was possible after having seen a strong picture. While this explanation is similar to the salience explanation, it assigns the locus of the effect to the pictures, and not the sentence interpretation, and to an awareness of a strong interpretation, and not the alternative.

To test the strong picture explanation we altered the false picture configuration for weak primes. Weak primes now consisted of a weak picture and a false strong picture, that is, a picture that had a strong configuration but was false by virtue of the sentence predicate. Thus all three primes involved a strong configuration picture.

Method

Participants

One hundred participants took part. Fifty participants were recruited online (via Prolific.ac.uk) and were paid, and fifty students were recruited for course credit from Cardiff University. Demographic information is only available for the online participants (29 male, average age 25.6 years). All stated they were native English speakers.

Design and Procedure

The design and procedure was the same as for Experiment 1 except for the construction of the weak primes. Weak primes consisted of a weak picture and a false picture, but unlike Experiment 1, the false picture had a strong configuration. The letters in the false picture were inconsistent with the sentence predicate (see Figure 5).

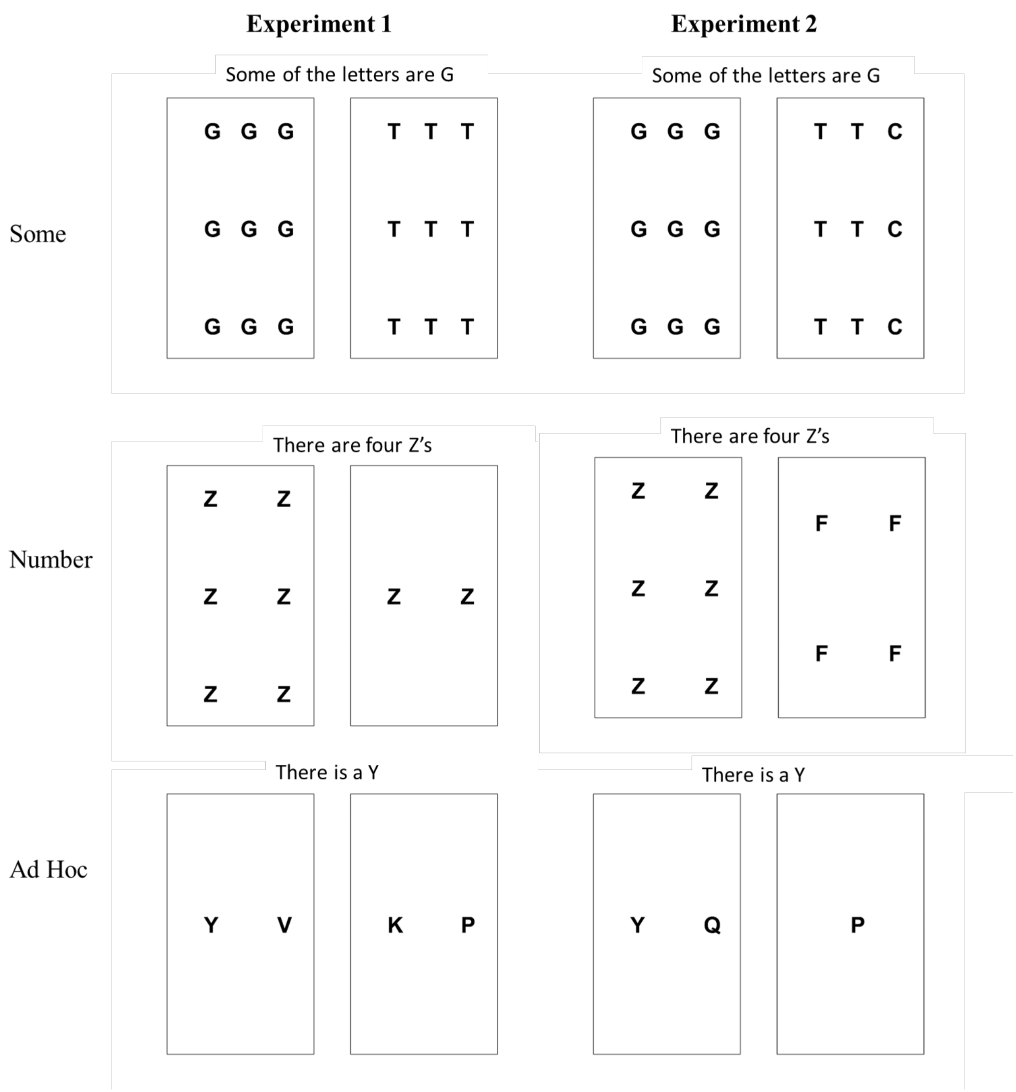

Figure 5. Change in weak stimuli from Experiment 1.

Results

Target trials with an incorrect preceding prime trial were removed from the analysis, as in Experiment 1. This accounted for 3% of the data. The remaining data underwent a logit transformation.

The main findings from Experiment 1 were replicated (see Figure 6). There was a main effect of prime type, $F(2, 184) = 49.56$ $p < .001$, and expression, $F(2, 184) = 39.58$, $p < .001$. The main effect of prime was caused by significantly lower implicature rates following the weak prime and strong/alternative primes, $t(99) = 7.60$, $p < .001$ and $t(99) = 8.35$, $p < .001$ respectively, combined with no difference between rates of implicature following strong or alternative primes, $t(99) = .13$, $p = .90$, BF = 0.11, in support of the salience model. The main effect of expression was driven by significantly lower rates of implicature for the *ad hoc* expressions compared to the *some* and number expressions. Unlike Experiment 1, there was a significant interaction between prime type and expression, $F(4, 368) = 4.24$, $p = .002$. We investigated this by testing the effect of the prime on each of the expressions.

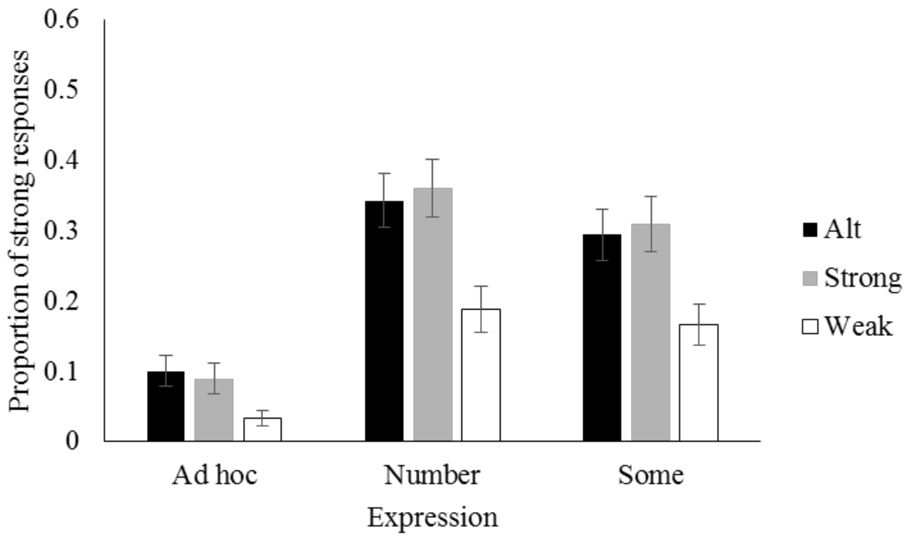

Figure 6. Proportion of strong (implicature) responses to target trials by expression. Error bars show standard error.

The interaction was driven by a significantly smaller effect of the strong/alternative prime on the *ad hoc* expressions compared to the *some* and number expressions. This is likely caused by floor effects for the *ad hoc* implicatures (overall, rates of implicature were lower in this experiment than in Experiment 1). When we removed *ad hoc* expressions from the analysis the interaction was no longer significant $F(4, 369) = 0.58$, $p = .56$, BF = 0.05.

More importantly, the same pattern of priming was found across all the expressions, just as in Experiment 1. There were significantly higher rates of implicature following strong and alternative primes compared to weak primes $t(94)$'s>2.66 p's <.009, $t(99)$'s>6.55, p's <.001, $t(97)$'s>4.92, p's <.006, and there was no difference between strong and alternative primes across expressions, t(94) = 1.34, $p = .18$, BF = 1.31; $t(99) = -.42$, $p = .68$, BF = 0.12; $t(97) = -.30$, $p = .77$, BF = .29, for *ad hoc*, number, and *some* expressions, respectively, in support of the salience model. Note that although the Bayes Factor for *ad hoc* expressions suggests that there is only limited support for the null hypothesis, the potential effect is in the opposite direction to the prediction of the combination model. The combination model predicts a greater rate of implicature following a strong prime than an alternative prime, the opposite of the results here.

Individual participant details

We conducted a similar type of individual participants analysis as that in Experiment 1. Figure 7 shows the results. Overall the pattern is very similar to Experiment 1 although there is an increase in the number of participants who maintained the same response throughout the experiment. Nonetheless, for *some* and the numbers around half of participants changed their response at least twice in the experiment, and for *ad hoc* expressions 15% changed their response at least twice. As in Experiment 1 our results cannot be explained by participants changing response only once during the experiment.

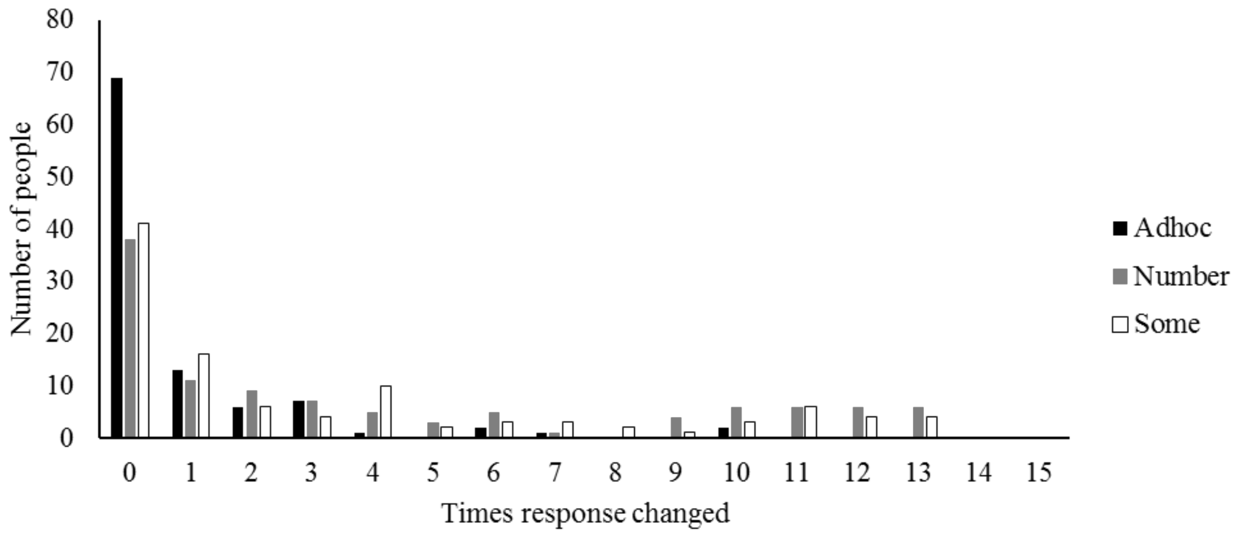

Figure 7. Frequency of response changes.

Summary

The priming effects we observed in Experiment 1 were replicated in Experiment 2. This provides more support for the salience model. Moreover, since the weak prime involved the same letter configurations as the strong prime, the priming effects cannot be explained solely by a visual priming effect.

Overall the rate of implicature was lower here than in Experiment 1, by around 10% (compare Figure 6 with Figure 3). This occurred across all three expression types. Could the cause of the lower rates of implicature be due to the change in the weak prime? This seems unlikely because in Experiment 2 there were more images corresponding to a strong interpretation than in Experiment 1. If anything, more strong images should have raised the implicature rate because participants would have been more aware of what a "better picture" might be. Consequently we attribute the difference across experiments to sampling variability (see Antoniou, Cummins, & Katsos, 2016, for a discussion of individual differences in scalar implicatures).

In Experiment 3 we test whether the priming effects might be due to a bias towards preserving a configuration->reference label mapping across trials.

Experiment 3

Consider the weak *some* prime in Figure 2. Here, the correct picture is made up of nine letters of one type (Z's). In the subsequent target trial the weak picture also has nine letters of the same type. Consequently, if the participant selects this picture they will have selected a picture with the same configuration (nine letters) and the same reference label (*some*) as the prime. In contrast, the correct response for the strong *some* prime has six letters of one type and three of another, but the weak picture in the target has nine letters of one type. If the participant selects the weak picture, they will have selected a card with a different configuration but the same label (*some*). This might feel incongruous because different pictures generally have different labels, and the participant might assume that since the label (*some*) is the same as the previous trial, the picture should also have the same configuration. Consequently they may prefer to select the "Better Picture" option. The strong-weak priming effect might therefore be explained by a desire to keep the configuration->reference label mapping consistent across trials. The alternative-weak priming effect can also be explained in this way. With the alternative *some* prime trial, the correct prime has nine letters, and the weak picture in the target trial also has nine letters. However, the label has changed between prime and target. In the prime trials, a picture with nine letters was described with the label *all* but in the target, a picture with nine letters was described with *some*. Thus, to maintain a consistent configuration->reference label mapping, the participant would select the "Better Picture" option. A similar form of reasoning can be used to explain the effects in the other expressions.

To test this hypothesis we altered the target trials so that instead of a scalar sentence, such as, "Some of the letters are A's," they contained an alternative sentence, such as, "All of the letters are A's" (see Figure 8). The target trials consisted of a weak picture, in which the letter configuration was consistent with the predicate (e.g., all nine letters were A's), and the "Better Picture" option. The prime trials remained the same as in Experiment 2.

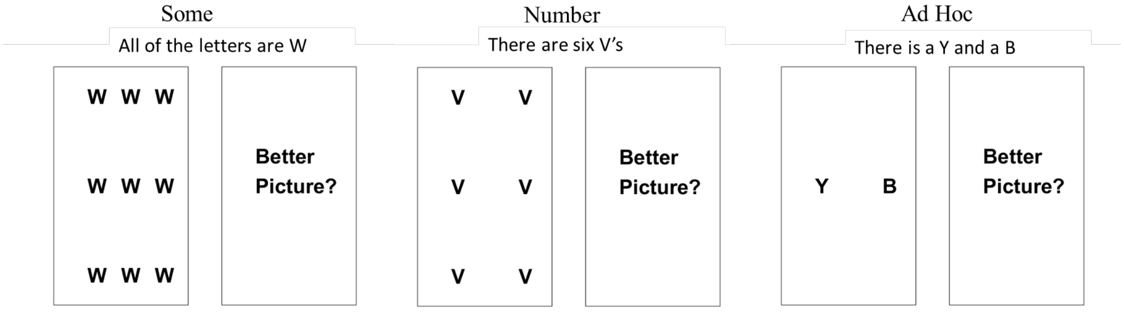

Figure 8. Experiment 3 Target trial configuration for some, number, and ad hoc expressions.

Consequently, when target trials were preceded by strong primes both the correct prime picture and the label were different between prime and target. Hence the weak picture in the target trial could be selected without breaking the configuration->reference label mapping. Similarly, when target trials were preceded by alternative primes, the weak picture could also be selected since configuration and label would be consistent across prime and target. However, when target trials were preceded by weak primes, selection of the weak picture in the target would be inconsistent with preserving the configuration->reference label mapping. Both prime response and weak target response have the same configuration but they would use different labels (see Figure 9). Hence in this case, a participant seeking to maintain a consistent mapping would choose the "Better Picture" option. The overall pattern of responding predicted by the configuration mapping hypothesis is that strong and alternative primes should lead to a low rate of "Better Picture" selections, while weak primes should lead to a high rate of "Better Picture" selections.

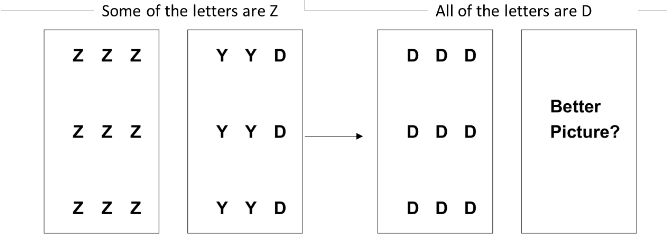

Figure 9. Example weak prime->target for Experiment 3.

Method

Participants

Twenty-five participants were recruited online from Prolific.ac.uk and were paid (15 male, average age 34.7 years).

Design and Procedure

The design and procedure were the same as in Experiment 2.

The sentence frame used in target trials was modified so that the sentence always used the alternative terms: *some* “All of the letters are [letter]”; number “Six of the letters are [letter]”; *ad hoc* “There is a [letter] and a [letter]”.

Thirty-six filler trials were included. Twelve of the filler trials had a similar construction as target trials but consisted of a false picture and a “Better Picture” option. In these trials participants should always select the “Better Picture” option. This was to ensure that participants considered the “Better Picture” as a valid choice in target trials. The remaining fillers were indistinguishable from prime trials. As in previous experiments, they served to break up the prime-prime-target structure.

Analysis and Discussion

The proportion of “Better Picture” selections following alternative, strong, and weak primes was 0.00, 0.01, and 0.02 respectively (for comparison, “Better Picture” selections to the false filler trials were at ceiling, 1.00). If there were a strong bias towards maintaining a consistent configuration-> reference label mapping across trials we would have observed a large increase in “Better Picture” selections following the weak prime. Since we did not, we can be confident the effects we observed in Experiments 1 and 2 were not due to a similar bias.

The configuration -> reference label hypothesis is part of a wider issue about visual similarity in priming tasks. Perhaps there is a general bias for people to select targets that are visually similar to the prime, irrespective of the sentence. In Experiments 1 and 2, the pattern of results we observed could be explained by an account along these lines. Participants chose the weak option in target trials after a weak and an alternative prime, which are structurally similar (e.g., in *some* trials, the weak prime has 9 identical letters, as does the weak target option), and the “Better picture” option after the strong prime, for which the weak target is not structurally similar (e.g., in *some* prime trials, the strong option has 6 letters of one type and three of another, but the weak option in the target does not). The results of Experiment 3 argue against this hypothesis, however. If visual similarity were the only explanation for the priming effect, strong primes should have been followed by a high rate of “Better Picture” responses (strong prime options differ substantially to the alternative target options) relative to weak and alternative primes (weak and alternative primes are structurally similar to the alternative target options), but these were not the results we observed. It might be possible to argue for a weak visual similarity account of our findings, however, in that the target cards were unambiguous in Experiment 3, and possibly the role of visual similarity is greatly weakened when the sentence is ambiguous. Our response is to note that the priming effects we observed in Experiments 1 and 2 were extremely large and if a visual bias were to exist, we would have expected to observe priming effects of some degree in Experiment 3, even if they were smaller effects, but we observed nothing at all. Furthermore, other researchers using extremely similar paradigms (but with ambiguous sentences) also conclude that visual similarity alone cannot explain their priming effects (see Bott & Chemla, 2014; Feiman & Snedeker, 2016; and Raffray & Pickering, 2010).

More generally, the results of Experiment 3 illustrate that the effects of the prime are restricted in ways that are theoretically sensible. The primes do not affect every target response. They only alter responses when there is a meaningful linguistic link between prime and target. Participants were not so influenced by a pragmatic need to please the experimenters that they altered their responses whenever they saw two consecutive primes, for example. Nor were they misunderstanding the role of the "Better Picture" option. They used the "Better Picture" option when they could think of a better picture that matched the meaning of the sentence, and not for other reasons.

General Discussion

This study tested between a salience model of implicature, in which the implicature decision is entirely determined by the activation of the alternative, and a combination model, in which an independent mechanism that uses the alternative also plays a role. While we found strong evidence that an active alternative primed implicatures, there was no evidence that the usage mechanism could also be primed. This pattern occurred in three types of scalar implicatures (*some*, numbers, and *ad hoc*) and across two experiments. We also conducted a third experiment to eliminate a possible relabelling explanation of the priming effect. Since none of our findings require the more complex, combination model, we favour the salience model.

An obvious concern is that our conclusion is based on the absence of an effect rather than an observed effect. There are two, connected, issues here. First, how can we be sure that our experiments were sufficiently powerful to detect a reasonably sized effect? And second, how do we know how large the usage priming effect should be? The answer to these questions lies partly with the Bayes Factor (BF) analysis that we used to test the null findings. The logic behind the BF is that the experimental hypothesis (in this case that there is priming of the usage mechanism) makes effect size predictions, as does the null hypothesis (that there is no usage priming), and the BF is the ratio of the likelihood of the experimental hypothesis explaining the data to the likelihood of the null hypothesis explaining the data. Under the effect size assumptions we made, and the data we observed, the odds were much higher in favour of the null hypothesis than the experimental hypothesis. Note that standard frequency-based statistics (p-values from e.g. t-tests) do not allow researchers to state evidence in favour of the null hypothesis (Rouder et al., 2009), as we do here, for a variety of reasons, such as the fact that p-values under the null hypothesis do not converge to a limit as sample size increases. With respect to the effect size, the analysis we report assumed an effect size distribution centred around a medium effect (a non-informative prior), but unless effect sizes of the usage priming mechanism are taken to be exceptionally small, the analysis always favoured the salience model.

A related issue is whether an independent usage mechanism exists but cannot be primed. This would be a mechanism that was affected by factors other than the salience of the alternative, but would not be sensitive to recent activity. Clearly, such a mechanism would not be observable with the paradigm used here. While we cannot eliminate this explanation, a large range of linguistic representations and processes have been successfully primed using similar paradigms (e.g., active versus passive forms, Bencini & Valian, 2008; transitive and dative syntactic forms, Bock 1986; animacy assignments, Bock, Loebell, & Morey, 1992; conceptual level structures for configurations in a maze, Garrod & Anderson, 1987; scopal interpretations with “every”, Raffray & Pickering, 2010; semantic coercion, Raffray, Pickering, Cai, & Branigan, 2013) and a non-primeable usage mechanism would be noteworthy for this reason alone. Furthermore, research on priming often has the tacit assumption that if the representation or mechanism of interest cannot be primed, no such representation exists (Branigan & Pickering, 2017; Pickering & Ferreira, 2008). For example, Raffray et al. (2013) suggest that a failure to find priming of noncompositional semantic-to-syntactic mappings would mean that there was no difference between compositional and noncompositional mappings. The difference between our study and those of other researchers is that we have drawn conclusions from negative observations rather than positive observations, but the logic of our study is identical. Finally, the priming effects shown in this study exclude the possibility that scalar implicature mechanisms in general are impervious to priming; the usage mechanism in particular would have to be non-primeable.

Priming scalar implicatures

We observed substantial priming of scalar implicatures. When the target trials were preceded by either the strong or the alternative prime trials, there were more implicatures relative to when the target was preceded by a weak prime. Whilst it is clear that scalar implicatures were primed, there are several potential explanations for why.

The first is that strong and alternative primes raised activation levels of the alternatives during target trials, whereas weak primes did not. Consequently more alternatives exceeded the threshold in the strong and alternative conditions, thereby resulting in more implicatures. This is the account we presented in the Introduction. The second is that participants were primed to view the context as one in which informativity, or precision, was important. Implicature derivation requires individuals to be sensitive to informativity. The listener must be aware that a speaker could have been more informative, i.e. could have used a stronger expression, and that the speaker was aware that they were less informative than they could have been. Accordingly the prime trials may have altered participants’ assumptions about acceptable levels of informativity (precision). The strong prime could have generated a context in which an informative, or precise, interpretation was expected. In target trials the expectation of informativity is best satisfied by the “Better Picture” option. Conversely the weak prime generated a loose, or “good enough” context. Consequently, in target trials, participants could have believed the weak image was good enough after the weak prime, since this satisfied the contextual demands. Alternative primes might also generate a precise context because the sentence was maximally informative, thus encouraging a more precise response to the subsequent target.

We cannot distinguish the two explanations from the data but the local priming effects we observed argue against the informativity account. We used a within-subject design in which the primes varied within the same experimental context. The individual participant analysis confirmed that participants switched between strong and weak interpretations of the same expression. Alternatives might be expected to vary in activation across short time intervals (i.e. locally) because linguistic contexts can change rapidly in dialogue (the same alternative might be active in one context but not in another a few seconds later). In contrast, precision is more likely to be a property of a global context, such as the style of individual speakers (e.g. X speaks precisely, Y does not), individual listeners (e.g. children might be tolerant of imprecision, and so are less likely to derive an implicature, relative to adults, as in Katsos & Bishop, 2011), or situational factors (e.g. one should be precise when describing experiments in reports, but loose is fine when texting).

A further explanation is that the priming effect was driven entirely by priming of the weak interpretation. Perhaps participants derived the strong interpretation by default, and were then reminded of the weak interpretation following the weak prime. There are two arguments against this explanation, however. The first is that responses were biased towards the weak interpretation, if anything, not towards the strong interpretation. This is particularly clear for the *ad hoc* expressions. Here, the overall proportion of strong responses was 0.15 in Experiment 1 and 0.08 in Experiment 2. If the default were indeed the strong interpretation, we would have expected much higher proportions of strong interpretations overall. The second is that participants repeatedly switched between strong and weak interpretations of the target within the task, as illustrated by the individual participants analysis. This means that strong and alternative primes must have shifted responses from weak to strong, at some point, rather than from strong to weak only. Participants responded weak after weak primes, and then strong after alternative/strong primes, then returned to weak after weak primes etc. The weak prime hypothesis could not explain the effect of switching from a weak response to a strong response.

The explanations above do not challenge the salience account. Whether activation, informativity, or default implicatures are the root cause of the priming effect, none require an independent usage mechanism. The final possibility we consider however, is a dual explanation model, which uses alternative activation and a usage mechanism. Consequently it supports a combination model. The dual explanation account proposes that the priming effect in the strong condition was caused by a combination of primed usage mechanism and primed alternatives, whereas that in the alternative condition was caused solely by the primed alternatives. The reason why the effect was the same size across conditions is that the alternative priming effect was greater in the alternative condition than the strong condition, which was then offset by priming of the usage mechanism. We argue against the dual explanation account, however, for the following reasons. First, it is more complex than the salience account but explains no more data. The dual account requires one set of factors to explain the alternative priming effect and another to explain the strong priming effect, and that these produce the same sized effects (in three expression types). In contrast, the salience account proposes one set of factors that explains both sets of effects, and the reason the effects are the same size is that they have the same underlying cause.

Second, the combination account assumes that the alternative priming effect is smaller in the strong condition than the alternative condition, and we see no reason for this assumption to be true. Note that the reason cannot be that the alternative was considered more often in the alternative condition because participants must have generated the alternative in the strong prime trials to select the correct response (and we omitted trials in which this did not happen). Thus the alternative must have been more active in the alternative condition, over and above that which caused the implicature to be triggered. One potential reason is that participants read the alternative in the alternative condition but derived it themselves in the strong condition. Perhaps comprehension is a more effective prime than production. But there is no evidence that comprehension-to-production priming (as in the alternative condition) is more effective than production-to-production priming (as in the strong condition). Certainly, structural priming studies demonstrate that comprehension to production priming exists (e.g. Branigan, Clelland and Pickering, 2000), but there are many production-to-production priming studies showing extremely large effects (e.g., Pickering & Branigan, 1998; see Pickering & Ferreira, 2008, for a review) and it is difficult to see why using comprehension procedures should prime production procedures more than production procedures themselves. A similar argument can be made from the education literature. Researchers have shown that re-reading text (comprehension) is not as effective a learning aid as self-generating answers (production; e.g., Dunlosky et al. 2013). One reason for this could be that production makes the information more salient than comprehension, contrary to the dual explanation account assumptions.

The dual explanation account of our findings is logically possible but we do not find any reason to reject the more parsimonious salience explanation. We found no difference between conditions using three expressions and we feel that the onus is now on proponents of the combination model to find positive evidence of an independent usage mechanism.

Alternatives in scalar implicatures

Our findings help to clarify and extend previous work on alternatives in scalar implicatures. Skordos and Papafragou (2016) found that when children were exposed to relevant alternatives, the rate of implicatures increased, but not when they were exposed to the same alternatives but in an irrelevant context. Our data suggests the salience model provides the best explanation for these findings. In the relevant condition of Skordos and Papafragou, it was necessary to process the quantifiers deeply in the training block to correctly match the picture to the sentence. Activation levels of the quantifiers were consequently high. Conversely, in the irrelevant condition, processing could be much more shallow, and quantifier activation levels were consequently low. In the subsequent testing block, high activation levels of the alternatives in the relevant condition translated into high rates of implicature. While this does not conflict with Skordos and Papafragou's claim that the deficit in children's processing lies in their understanding of relevance (as compared to a deficit in lexical retrieval of the alterative), it shifts the locus of the explanation away from how alternatives are used by children towards how activation levels are computed, and how this might differ between adults and children.

Our study also extends Skordos and Papafragou (2016) by demonstrating that alternative activation is a determiner of scalar implicatures in numbers and *ad hoc* scales, not just *some*. This is important because it suggests that the process by which strong and weak meanings are related in numbers and *ad hoc* implicatures is the same as in *some*. While *some* is the prototypical scalar implicature, there is some controversy about whether the strong number meaning (*exactly* N) is derived from a lexically specified weak number meaning (*at least N*), as in standard scalar implicatures (e.g., Horn, 1972; 1989), or whether the weak meaning is derived from a lexically specified strong meaning (e.g., Breheny, 2008, Geurts, 2006; and see Huang, Spelke & Snedeker, 2013, for psychological evidence). Since strong interpretations can be primed in similar ways for numbers and *some*, the most straightforward conclusion is that the derivation and representation processes are also similar.

This is not to say that there were no differences between the different expressions in our task. While priming effects were similar across expressions, *ad hoc* expressions had a much lower rate of implicature than *some* or the numbers. One explanation is that the alternative for *ad hoc* expressions was more difficult to derive than the alternative for the other expressions (see Bott & Chemla, 2016, for a similar point). For *ad hoc* expressions, the alternative required adding material to the sentence (e.g. the alternative to "There is an A," was, "There is an A and a D") whereas for *some* and the numbers, the alternatives involved replacing an element of the sentence with another element (e.g. the alternative to "Some of the letters are As" was "All of the letters are As"). Expansion could be more complex for the language processor than replacement. This mirrors Katzir (2007), who argues that there is a fundamental difference between alternatives formed from replacement and those formed by the addition of material.

There are some problems with this explanation, however. One is that in the alternative prime condition, participants were presented with the alternative structure immediately prior to the target. The difficulty involved in generating an alternative for the *ad hoc* expressions must have been minimal in this condition, yet differences across expressions were large. Another is that the developmental data isn't consistent with the claim that *ad hoc* alternatives are more difficult to derive than *some* alternatives (*all*). Barner, Bale and Brooks (2010) showed that while 4 year old children had difficulty retrieving the alternative for *some* (as did Skordos & Papafragou, 2016), they easily constructed the alternative for context dependent conditions, which were similar to the *ad hoc* expressions used here. If the alternative for *some* is easy to derive, one might expect it to be acquired early. Consequently, it might be that complexity plays no part in the construction of alternatives after all, and instead, the *ad hoc* alternatives are less salient for some more general contextual reason.

Another possibility is that the difference in rate of implicatures across expressions is due to differences in the baseline activation of an independent usage mechanism. Participants might have had less experience with generating *ad hoc* implicatures than quantifiers and numbers, which in turn would lower the baseline activation for the usage mechanism in these cases. We cannot rule this explanation out completely but we make two arguments against it. First, this explanation requires a separate usage mechanism for each expression (a separate usage mechanism for *ad hoc* implicatures, quantifiers, and numerals). Accounting for the difference in implicature rates on the basis of the salience of the alternatives requires only a single usage mechanism for all three expressions. The salience explanation is therefore more parsimonious. Second, if the different usage mechanisms obtain different levels of baseline activation through experience, it is not clear why they could not be primed in our experiment. Developmental priming of structures is generally assumed to operate using the same mechanism as short term priming (Pickering & Ferreira, 2008) and in which case, a structure that is susceptible to developmental priming should also exhibit short term priming.

Our results also bear on the interpretation of Bott and Chemla (2016). Recall that Bott and Chemla observed priming of implicatures within the same expression (e.g. from *some* to *some*), just as we did. However, they also found that different sorts of implicatures primed each other (e.g., *some* primed numbers). They suggested that this effect could be explained in two ways. The first assumes that a usage mechanism was primed. The strong prime trials elevated the activation of the usage mechanism so that when reading the ambiguous target trials, participants were more likely to use the alternative, i.e. combine it with the literal meaning. This explanation is consistent with a combination account, in which alternative salience and alternative usage make different contributions to the rate at which people derive implicatures. The second was that an alternative search mechanism was primed. In the former, a mechanism could be triggered to search for alternatives so that the mechanism was more active after a strong prime trial than a weak prime trial, regardless of whether the expressions were of the same type. This explanation is compatible with the salience model. Since we found no evidence for the primeable usage mechanism, our data suggest that the search mechanism was the more likely explanation for Bott and Chemla's between expression priming results.

Finally, Katsos and Bishop (2011) found that in some contexts, children were aware of alternatives but chose not to derive the implicature (they were more "tolerant" of underinformative sentences than adults). This finding is difficult to reconcile with our conclusion: If there is no independent usage mechanism, how can people be aware of the alternative but not derive an implicature? We are not confident of the answer but we suggest the difference might lie in a cancellation mechanism. While scalar implicatures might always be derived given sufficient salience of the alternative, there might be a cancellation mechanism that applies after the fact, as in traditional defeasement. It might be this mechanism that differs across adults and children, not the initial derivation of the implicature.

Scalar implicatures and representations

Scalar implicatures are typically described in terms of mechanisms, such as deriving the alternatives and adherence to maxims, in keeping with pragmatics tradition, whereas structural priming researchers refer to representations, in keeping with syntax and semantics. However, Bott and Chemla (2016; see also Rees & Bott, 2017) suggest that implicatures could also be seen as meaning-based representations that can be primed (similar to Raffray and Pickering's (2010) account of scopal ambiguity priming). They note that the processor could access a representation for scalar implicature, [S + not(Alt(S))], where S corresponds to the sentence and Alt a more informative sentence (an alternative), and that this representation could be linked to trigger expressions such as *some*. Both the link between the trigger expression and the representation itself could be primed, much like the representations and links assumed in the structural priming literature (e.g. between verbs, such as *give*, and syntactic frames, such as double object constructions, see Pickering & Branigan, 1998). We agree that it is useful to consider implications in this way but our results argue against the equivalence of the representations proposed in the structural priming literature and an [S + not(Alt(S))] implicature representation, as we describe below.

That there are meaning-based representations associated with scalar implicatures is indisputable. Since implicatures are part of what is communicated, implicatures must be represented by the listener. However, out results demonstrate that this representation is not independent from the material that gave rise to it, in the sense proposed by researchers in the structural priming literature. It is not a static, pre-existing frame that is waiting to be filled by appropriate content, as it is in a syntactic frame. Two aspects of our finding give rise to this conclusion. First, if there were such a representation, we should have been able to prime it. We used the same method that has been used to discover independent representations in the structural priming literature yet we saw no evidence of a primeable implicature representation. Second, we found that alternatives primed scalar implicatures. Although this does not conflict with the existence of an [S + not(Alt(S))] representation, there is no evidence that structural priming representations, such as syntactic frames, are influenced by an alternative (indeed, syntactic frames do not have alternatives in the same way that utterances do). Overall, our work suggests that implicature representations are quite different to those proposed in the structural priming literature (c.f. Rees & Bott, 2017).

Default scalar implicatures and salience

The salience account is similar in many respects to the Neo-Gricean, default model (e.g., Levinson, 2000; Horn, 1972). In both cases, the implicature is obligatorily derived (i.e., by default) providing the alternative is salient. How then, does the salience account explain the processing cost associated with scalar implicatures, as observed by Bott and Noveck (2004) and others (e.g., Breheny, Katsos & Williams, 2006; Huang & Snedeker, 2009)? The default model, as expressed by Bott and Noveck, stated that the implicatures are always derived, but that they can sometimes be cancelled. Thus there is no explanation for why the weak interpretation of a sentence can ever be faster to understand than the strong interpretation, as per their results. Under the salience account, however, the strong interpretation will only arise if the alternative is sufficiently salient. If the context does not sufficiently raise the salience of the alternative, the strong interpretation will not arise. Thus, in the literal conditions of Bott and Noveck, the argument would be that the alternatives were not sufficiently salient, and so no implicature was derived. Conversely, in the implicature conditions, the alternative was made salient, and so the alternatives were negated, incorporated into the sentence etc., with the accompanying processing cost. Politzer-Ahles and Gwilliams (2015) and Degen and Tanenhaus (2015) make similar arguments. Of course, this is only a hypothesis, and we look forward to further work testing whether processing costs associated with scalar implicatures are reduced when alternatives are salient.

Conclusion

The aim of these studies was to investigate the role that alternatives play in the derivation of implicatures. Previous work has focussed on how the alternatives are used, whereas here we focussed on the role played by salience. Our study makes two important contributions. (1) We demonstrate that adults are sensitive to the salience of alternatives when deriving scalar implicatures. This is a local effect, not dependent on using different speakers or contexts to vary salience, and applies across a wide range of purported scalar implicatures. (2) We find no evidence of a usage mechanism that applies independently of the salience of the alternative. We thus suggest that the rate of scalar implicature is determined entirely by the salience of the alternative.

Supplementary Material

All data for the experiments presented can be found on the Open Science Framework https://osf.io/986c2/#.

References

Antoniou, K., Cummins, C., & Katsos, N. (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics*, *99*, 78-95.

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition, 118*, 84–93.

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language, 59*, 97-113.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, J. K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*, 437-458.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150-171.

Bott, L., Bailey, T., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language, 66*, 123-142

Bott, L., & Chemla, E. (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language, 91,* 117-140.

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*, 437–457.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co- ordination in dialogue. *Cognition, 75*, 13–25.

Branigan, H. P., Pickering, M. J., Liversedge, S.P., Stewart, A. J., & Urbach, T. P. (1995). Syntactic priming: investigating the mental representation of language. *Journal of Psycholinguistic Research, 24,* 489-506.

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25*, 93–140.

Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition*, *126*(3), 423-440.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*, 434–463.

Chemla, E., & Bott, L. (2014). Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice. *Cognition, 130*, 380-396.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention* (Vol. 2). Oxford University Press.

Chierchia, G., Crain, S., Gausti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. BUCLD 25 Proceedings.

Chierchia, G., Fox, D., & Spector, B. (2012). Scalar implicature as a grammatical phenomenon. In P. Portner, C. Maienborn, & K. von Heusinger (Eds.). *An international handbook of natural language meaning* (Vol. 3, pp. 2297–2331). Mouton de Gruyter.

Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive Science, 39*, 667–710.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? Perspectives on *Psychological Science, 6*, 274–290.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 1–17.

Feiman, R. & Snedeker, J. (2016). The logic in language: How *all* quantifiers are alike, but *each* quantifier is different. *Cognitive Psychology, 87,* 29-52.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition, 27*, 181-218.

Geurts, B. (2006). Take ‘‘five”: The meaning and use of a number word. In V. Svetlana & L. Tasmowski (Eds.), *Non-definiteness and plurality* (pp. 311–329). Amsterdam: Benjamins.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Gotzner, N., Wartenburger, I., & Spalek, K. (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition, 8*, 59-95.

Grice, H. P. (1989). *Studies in the way of words*. Massachusetts: Harvard University Press.

Grodner, D., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). ‘‘Some”, and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition, 116*, 42–55.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Gausti, M. T. (2001). At the semantics/pragmatics interface in child language. Proceedings of the 11th semantics and linguistic theory conference.

Hirschberg, J. (1991). *A theory of scalar implicature*. New York: Garland Press.

Horn, L. R. (1972). On the semantic properties of logical operators in English. Ph.D. thesis, University of California, Los Angeles.

Horn, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Huang, Y. T., & Snedeker, J. (2009a). On-line interpretation of scalar quantifiers: Insight into the semantic-pragmatics interface. *Cognitive Psychology, 58*, 376–415.

Huang, Y. T., & Snedeker, J. (2009b). Semantic meaning and pragmatic interpretation in 5-year olds: Evidence from real-time spoken language comprehension. *Developmental Psychology, 45*, 1723-1739.

Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What exactly do numbers mean? *Language Learning and Development, 9,* 105–129.

JASP Team (2016). JASP (Version 0.7.5.6) [Computer software].

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition, 120*, 67-81.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy, 30*, 669–690.

Kim, C.S., Gunlogson, C., Tanenhaus, M.K., & Runner, J.T. (2015). Context-driven expectations about focus alternatives. *Cognition, 139*, 28-49.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Noveck, I. A. (2001). When children are more logical than adults: experimental investigations of scalar implicature. *Cognition, 78*, 165-188.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition, 86,* 253–282.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*, 427.

Politzer-Ahles, S. & Gwilliams, L. (2015). Involvement of prefontal cortex in scalar implicatures: evidence from magnetoencephalography. *Language, Cognition and Neuroscience, 30,* 1-14.

Raffray, C. & Pickering, M.J. (2010). How do people construct logical form during language comprehension? *Psychological Science, 21*, 1090-197.

Raffray, C. N., Pickering, M. J., Cai, Z. G., & Branigan, H. P. (2014). The production of coerced expressions: Evidence from priming. *Journal of Memory and Language, 74*, 91–106.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy, 27*, 367–391.

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition, 153*, 6-18.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition, 108*, 51–68.

Tomlinson, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language, 69*, 18-35.

van Rooij, R. & K. Schulz (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy 29,* 205–250.

van Tiel, B., & Schaeken, W. (2016). Processing conversational implicatures: Alternatives and counterfactual reasoning. *Cognitive Science*.